Colon carbonic anhydrase 1: transactivation of gene expression by the homeodomain protein Cdx2

F.-J. Drummond, J. Sowden, K. Morrison, Y.H. Edwards*

MRC Human Biochemical Genetics Unit, University College London, Wolfson House, 4, Stephenson Way, London NW1 2HE, UK

Received 12 December 1997

Abstract The homeodomain protein, Cdx2, has been implicated in the transcriptional regulation of genes expressed in the small intestine. In vitro studies of the carbonic anhydrase 1 (*CA1*) colon promoter implied that Cdx2 may also play a role in the regulation of colon-specific gene expression. The current work follows up this proposal by examining the ability of Cdx2 to transactivate gene expression in cultured cells mediated by *CA1* promoter sequences. The results show that Cdx2 exerts a positive regulatory effect by binding to a motif 87 bp upstream of the *CA1* TATA box; this motif appears to act as an enhancer since gene activation is independent of its orientation.

Key words: Intestinal gene regulation; Carbonic anhydrase; Cdx2; Caudal-type homeobox protein; Colon promoter

1. Introduction

Carbonic anhydrase (CA) is located in the cytoplasm and apical cell membrane of colon epithelial cells. The colon epithelium conducts the absorption and secretion of ions and water via its extensive and accessible cellular surface. The function of CA in the colon is principally that of electroneutral sodium chloride reabsorption and short chain fatty acid (SCFA) uptake [1,2]. CA catalyses the hydration of CO_2 to form HCO_3^- and H^+ ions which are subsequently exported from the cell in exchange with Na^+ and Cl^- ions [1]. The secretion of HCO_3^- is also important for alkalinisation of the luminal contents thereby maintaining a favourable pH for bacterial flora, and for active exchange for SCFAs [2]. H^+ ions secreted into the lumen of the colon are also involved in the passive, non-ionic, diffusive uptake of SCFAs [3].

Three carbonic anhydrase genes are expressed in the colon, *CA1*, *CA2* and *CA4* [4–7]. CA1 and CA2 proteins are found in the cytoplasm of epithelial cells, while CA4 is membrane-bound. CA2 and CA4 are found throughout the length of the gastrointestinal tract whereas CA1 is specifically expressed at high levels in the colon [8,9].

The *CA1* gene is unusual amongst the carbonic anhydrases in having two cell-type-specific promoters which act in a mutually exclusive manner. The proximal promoter transcribes *CA1* in colon epithelia, while the more distal promoter is active in erythroid cells [9,10]. While the regulation of the *CA1* gene in erythroid cells has been well characterised [10,11], the factors which control the expression of the *CA1* gene in intestinal cells are not known.

Our earlier studies using electrophoretic mobility shift assays (EMSA) identified a region within the proximal colon promoter of the human *CA1* gene which binds a factor present only in colon cells [12]. Sequence comparisons showed that the DNA motif binding this factor was similar to an element in the promoter of the small intestinal gene, sucrase-isomaltase (*SI*), which binds the homeodomain protein, Cdx2 [13]. Cdx2 has been shown thus far to activate the expression of a number of small intestine genes [13–15], although Cdx2 is abundant both in the small intestine and colon [16,17]. In the present study we have sought to confirm that Cdx2 is interacting with the *CA1* colon promoter sequences in a biological system, and in a manner which leads to either activation or repression of gene transcription. This question was examined directly by co-transfection of recombinant mouse *Cdx2* together with DNA constructs comprising a minimal promoter downstream of various *CA1* sequences, into cells which normally do not express either *CA1* or *Cdx2*.

2. Materials and methods

2.1. Preparation of CA1 promoter/reporter gene constructs

Regions of the human *CA1* colon promoter (Fig. 1) were cloned into the pCAT3-promoter vector (pSV40/CAT, Promega) upstream of the SV40 early promoter. These included a double-stranded 17 bp oligonucleotide −118 bp to −102 bp of the *CA1* promoter encompassing the core motif -TTTTACAA- (CPO17); an 89 bp PCR fragment −147 bp to −59 bp (CP1) extending across the same motif (Fig. 1) and a 166 bp PCR fragment comprising −224 bp to −59 bp (designated CP1/2) containing the CPO17 motif and a second putative Cdx2 binding element (Fig. 1). The antisense CPO17 oligonucleotide was given ends compatible with *Sac*I and *Xho*I overhangs, to allow directional cloning into pSV40/CAT. The CP1 and CP1/2 PCR fragments were inserted, in their natural and reverse orientations by TA-cloning into T-tailed *Sma*I-cut pSV40/CAT. Diagnostic digests were performed to confirm the identity and orientations of the inserts. Constructs were sequenced using an ALFexpress automated DNA sequencer (Pharmacia). For high levels of Cdx2 expression, the plasmid pRc/CMV-Cdx2 (pCdx2) containing the *Cdx2* cDNA downstream of a CMV promoter was used.

2.2. Transfection studies

HeLa cells for transfection experiments were cultured to 60–70% confluence, as previously described [12]. Cells were seeded at 5×10^5 cells in 60 mm dishes 24 h prior to transfection. 5 μg of each *CA1* reporter construct was co-transfected with either 5 μg of pCdx2 or 5 μg of pSV-β-Gal. Transfections were carried out using the calcium phosphate ($CaPO_4$) procedure exactly as described [18] except that cells were incubated for 20 h with the $CaPO_4$/DNA precipitate, prior to washing and further culture for 28 h. Transfections were tested in two identical dishes of cells and each experiment was repeated at least three times (in some cases 10 times, all in duplicate). CAT activity was monitored by two methods, thin layer chromatography (TLC) and liquid scintillation counting (LSC). Cell extracts were heat treated at 60°C for 10 min prior to assaying. Plasmids used as controls in transfection experiments included pSV40/CAT to estimate basal levels of CAT expression and pCAT control (Promega) to provide high levels of CAT expression. Purified CAT enzyme (Boehringer) was also used as a control in TLC assays.

*Corresponding author. Fax: (44) (171) 387 3496.
E-mail: yedwards@hgmp.mrc.ac.uk

PII S0014-5793(98)00103-3

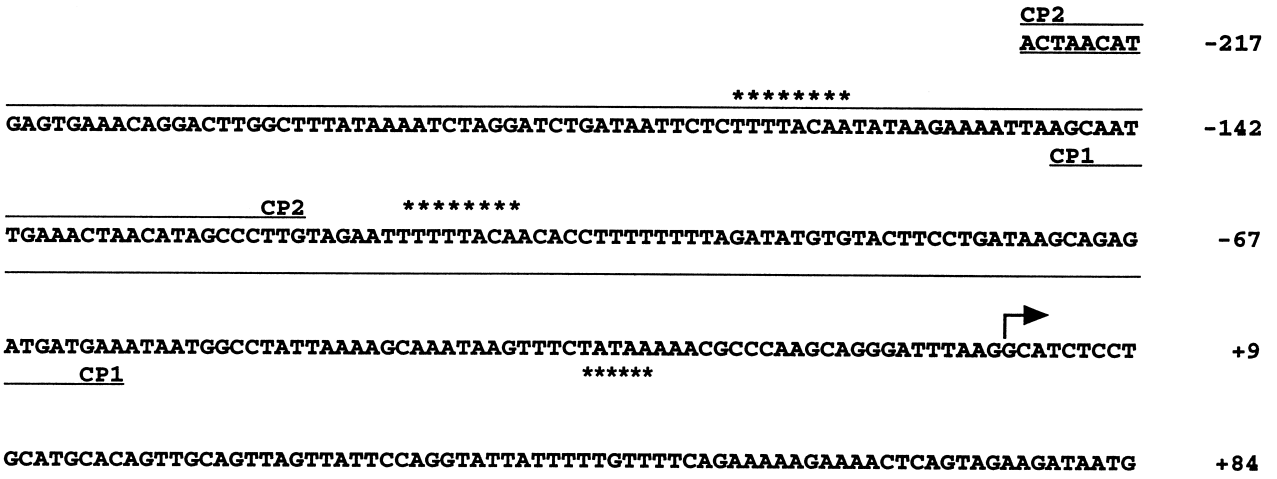


Fig. 1. Sequence of the human proximal *CA1* colon promoter. The −59 to −147 bp fragment (CP1) is indicated by underlining and the −122 to −224 bp fragment (CP2) by overlining. The transcription start site +1 is indicated (angled arrow). Asterisks indicate the position of two potential Cdx2 binding sites and the TATA box.

2.3. TLC assays

For TLC assays 100 μl of cell lysate was mixed with 3 μl of [^{14}C]chloramphenicol (0.15 μCi) and 5 μl of *n*-butyryl CoA, the volume adjusted to 125 μl with distilled water and incubated at 37°C for 3 h. Samples were extracted with 500 μl ethyl acetate; the organic phase was dried and resuspended in 20 μl of ethyl acetate for spotting onto TLC plates (Merck). Unmodified substrate and C1- and C3-butyrylated chloramphenicol were resolved by TLC in 250 ml of chloroform:methanol (97:3) at room temperature for 1 h. Plates were air-dried and autoradiographed with Kodak X-ray film.

2.4. LSC assays

For LSC assays sample reactions were carried out as for TLC assays except that butyrylated products were extracted with 300 μl of xylene. The organic phase was further extracted with 100 μl of 0.25 M Tris-HCl and 150 μl of the organic phase added directly to 1 ml of liquid scintillant and counted in an LKB 1211 minibeta scintillation counter. The value obtained for untransfected HeLa cells was deducted from that obtained for transfected cells. Duplicate estimates were averaged and expressed as a percentage of those obtained with the pCAT control vector transfected in the same experiment.

2.5. Electrophoretic mobility shift assays

Preparation of whole cell extracts from the colon carcinoma cell line HT115, labelling of double-stranded oligonucleotides and EMSAs were carried out as described previously [12]. Competitor oligonucleotide was added to the DNA/protein binding mix prior to incubation.

3. Results

3.1. Analysis of gene activation mediated by Cdx2 and CA1 promoter sequences

Our earlier DNase I hypersensitive site mapping and EMSAs identified an element in the *CA1* promoter which binds a protein found only in cells derived from the colon (shown as COF1/1′ in Fig. 6A) [12]. The protein binding element was contained within an 89 bp stretch of sequence (−147 to −59 bp, designated CP1) (Fig. 1); after making various deletions of this sequence it was found that binding activity was retained by a 17 bp sequence (−118 to −102 bp, designated CPO17). Sequence comparisons showed that the core sequence of this motif, -TTTTACAA-, was similar to elements found in the promoters of the small intestinal genes, *SI* and lactase (LPH) which are known to bind the homeodomain protein, Cdx2 [13,15].

Since Cdx2 is expressed both in small intestine and colon it seems reasonable to infer that this transcription factor also plays a role in the regulation of genes in the colon in vivo. We have investigated this possibility by assessing whether Cdx2 can modify gene transcription in cultured cells, by interaction with *CA1* sequence motifs. Transient transfection

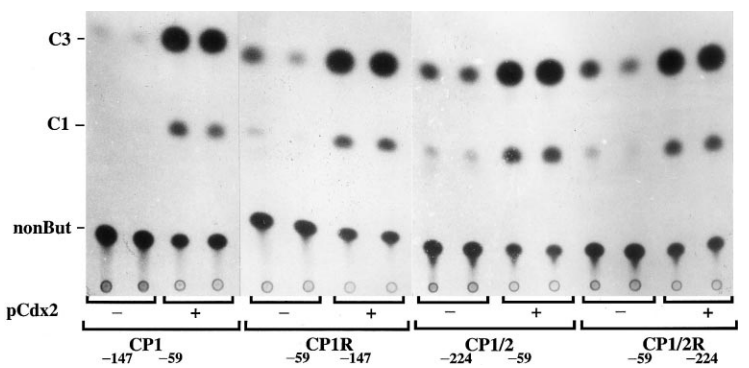


Fig. 2. TLC assay of duplicate transfections of HeLa cells with *CA1*/SV40/CAT constructs in the presence (+) or absence (−) of pCdx2. The position of each *CA1* promoter fragment relative to the transcription start site is indicated and those in reverse orientation are shown as R. C1 and C3 indicate the butyrylated products of chloramphenicol; nonBut indicates unmodified substrate.

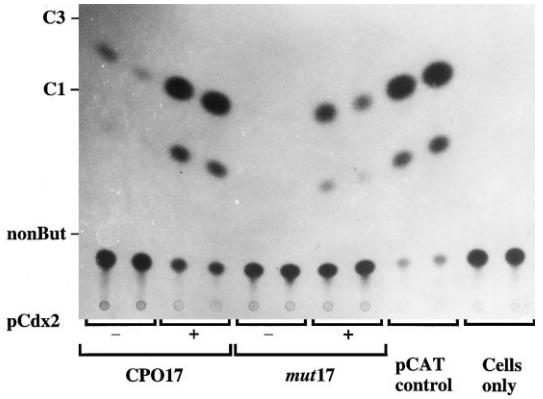


Fig. 3. TLC assay after duplicate transfections of HeLa cells with constructs containing *CA1* −118 bp to −102 bp (CPO17) and its mutated version (*mut*17), in the presence (+) or absence (−) of pCdx2, or with the positive control plasmid, pCAT control. Untransfected cells are also shown. C1 and C3 indicate the butyrylated products of chloramphenicol; nonBut indicates unmodified substrate.

assays were performed using various fragments of *CA1* promoter sequence (but excluding the TATA box), attached to an SV40 minimal promoter and upstream of the CAT reporter gene. HeLa cells were chosen because they do not normally express either *CDX2* (or *CA1*) and complications due to the presence of endogenous gene activity are not an issue. Changes in the level of CAT gene transcription were judged from measurements of CAT enzyme activity using thin layer chromatography (for examples see Figs. 2 and 3) and liquid scintillation counting (data summarised in Fig. 4).

The *CA1* DNA promoter fragments contained in the constructs used in these studies were as follows: −118 bp to −102 bp of the *CA1* promoter encompassing the single copy of the core motif -TTTTACAA- (CPO17); −147 bp to −59 bp (CP1), which contains this motif and 80 bp of additional sequence; and −224 bp to −59 bp (designated CP1/2) which also contains further 5′ sequences including a second copy of the core motif -TTTTACAA-, identified 54 bp upstream (Fig. 1). Co-transfection with a Cdx2 expression vector led to an increase in CAT gene activity by all constructs (Figs. 2 and 3). Overall, the highest level of activity, 64% of the pCAT-control vector, was found for the shortest fragment, −118 bp to −102 bp. In general the constructs containing longer CA1 promoter elements showed relatively lower levels of gene activation, 36% and 44% activity were found for the −147 bp to −59 bp (CP1-CAT) and −224 bp to −59 bp (CP1/2-CAT) respectively (Figs. 2–4). The reason for this is not clear but may reflect some interaction between the *CA1* sequences and other factors in the HeLa cell extracts which interfere with transcription of the reporter gene. Interestingly the −147 bp to −59 bp and −224 bp to −59 bp fragments of *CA1* promoter sequence showed 41% and 39% of pCAT-control activity respectively, when placed in a reverse direction relative to the SV40 promoter (Figs. 2–4). In Fig. 4, the differences between the means, with and without pCdx2, were tested for significance using Student's *t*-test, and in all cases the increase in CAT activity was statistically significant except for the pCP1-CAT which was borderline, $P = 0.1$.

In the *SI* gene two copies of the Cdx2 binding motif occur as an inverted repeat [13]; in the lactase gene only a single copy is found [15]. Suh and colleagues [13] have shown that mutation of the central core of the SI binding site, GCAATAAAACTTTATGA to GCAAGCCAACGGGATGA abolishes gene activation by the SI promoter. In order to confirm the importance of the *CA1* promoter motif in Cdx2 binding, the binding site (-AATTTTTTACAACACCT-) was mutated to change the central core from -TTTTACA- to -TTTGCCA-

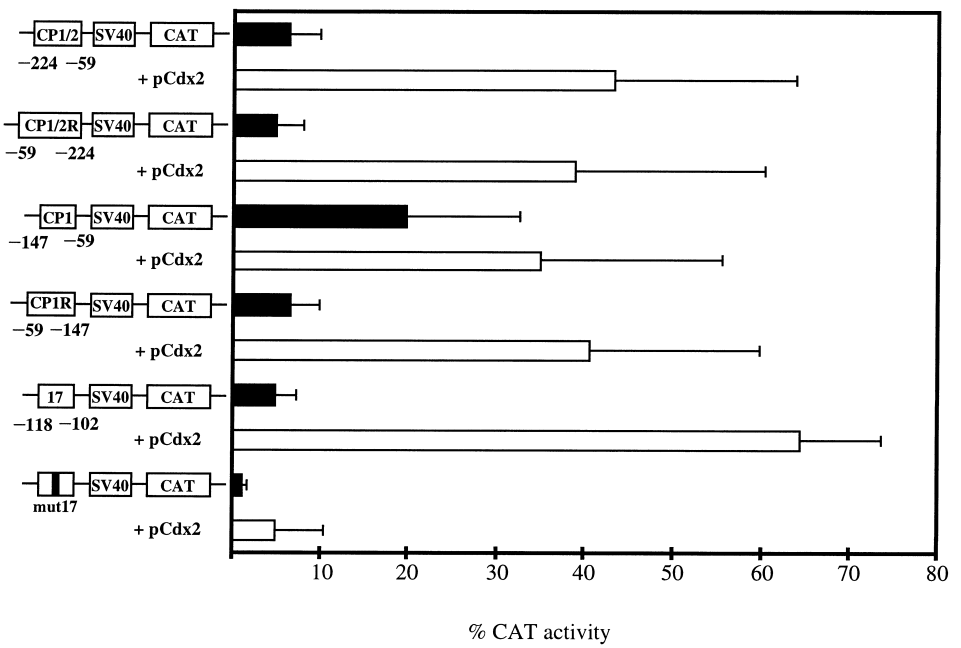


Fig. 4. CAT assays after transient transfection of *CA1*/SV40/CAT constructs in HeLa cells. [^{14}C]Butyrylated chloramphenicol was assayed by liquid scintillation counting in duplicate and the results from three to ten experiments were pooled and averaged. The constructs are shown to the left; the position of each *CA1* promoter fragment relative to the transcription start site is indicated. Those fragments in reverse orientation are indicated as R. Constructs were transfected alone (solid bars), or together with pCdx2 (open bars). Results for each construct are expressed as a percentage of the pCAT control activity ± S.D.

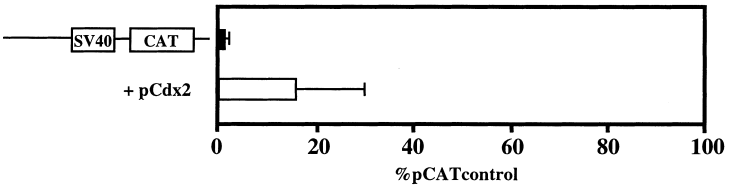


Fig. 5. LSC assays showing CAT gene activation mediated by Cdx2 binding to the SV40 minimal promoter.

(Figs. 3 and 6B). The ability of this construct to mediate gene activation was much reduced such that CAT activity was only 5% of pCAT control in the presence of Cdx2, reflecting a marked decrease in the ability to bind Cdx2 at the mutated sequence (Figs. 3 and 4).

Several observations can be made from these results; Cdx2 appears to exert a positive regulatory effect by binding to a *CA1* sequence element, -AATTTTTTACAACACCT-; the *CA1* sequence element appears to act as an enhancer element since gene activation induced by Cdx2 is independent of the orientation of the *CA1* binding motif; and the presence of two copies of the Cdx2 binding motif (as in the −224 bp to −59 bp construct) does not increase gene activity above that seen in the presence of a single copy of the core motif -TTTTACA- alone.

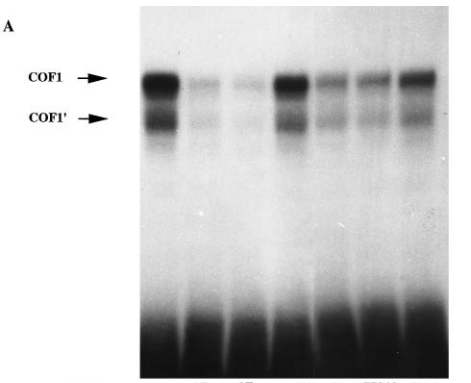
3.2. Activation of the SV40 minimal promoter by Cdx2

The estimates of relative CAT gene activity summarised in Fig. 4 were made after deduction of background activities due to the SV40 minimal promoter, which were assessed separately for each set of transfections. It was noted, incidentally, that transcription from the SV40 minimal promoter alone was increased to a low, but significant level (16% of pCAT control) in the presence of Cdx2 (Fig. 5). We considered the possibility that this activation was due to Cdx2 binding to a motif within the SV40 minimal promoter. Sequence comparisons pointed to the SV40 TATA box, which has the core sequence -TTTATG-, identical to the core Cdx2 binding motif in the SI gene promoter [13] and indeed an oligonucleotide encompassing the SV40 TATA box when used as competitor in competitive binding assays, partially competed for the binding of Cdx2 to the *CA1* −118 bp to −102 bp (CPO17) element (Fig. 6A). This finding is similar to that of Lambert and colleagues [19] who demonstrated Cdx2 binding to the TATA box of the rat calbindin-D9k (*CaBP9k*) gene, and raises the possibility that Cdx2 might perhaps bind to the TATA box of the human *CA1* gene itself. Both *CA1* and human *CaBP9k* TATA box motifs were tested in competition assays and both partially competed for the binding of the *CA1* −118 bp to −102 bp (CPO17) element, but the *CA1* TATA motif, in particular, competed much less efficiently than the *CA1* −118 bp to −102 bp (CPO17) element (Fig. 6A).

4. Discussion

Evidence from the studies presented here, taken together with our earlier EMSA and supershift assays [12], leads to the conclusion that the homeodomain transcription factor, Cdx2, is important in the regulation of colon-specific carbonic anhydrase 1 (*CA1*). The distributions of colon CA1 and Cdx2 are consistent with a role for Cdx2 in the regulation of *CA1* in vivo, since both are confined to the intestine epithelium and are most abundant in proximal colon; however, *CA1* expression is limited to the colon (large intestine) whereas *Cdx2* is also expressed in the small intestine [8,16]. Cdx2 has recently been recognised as an important transcription factor in the regulation of a number of genes expressed in the small intestine – sucrase-isomaltase (*SI*) [13], lactase-phlorizin hydrolase (*LPH*) [15], intestinal phospholipase A/lysophospholipase (*IPAL*) [14], calbindin 9K (*CaBP9k*) [19] – and pancreatic islet cell expressed genes – proglucagon [20] and insulin [21]. Furthermore Cdx2 has been shown to induce morphological and molecular differentiation in undifferentiated intestinal cell lines [22]. Recently Taylor and colleagues [23] have demonstrated that Cdx2 is able to activate transcription from the *SI* promoter and from heterologous promoters when the *SI* Cdx2 binding element is placed in an enhancer context. Our finding that gene activation mediated by a Cdx2 binding element in the *CA1* promoter is orientation independent is in line with this observation.

It is relevant to note that the *Cdx2* null heterozygote mouse is characterised by the presence of multiple colonic tumours which occur most frequently in the proximal portion of the colon [24] where CA1 is most abundant [6]. As yet there is no information regarding expression of *CA1* in these mice but



B

5'-AATTTTTTACAACACCT-3'	CPO17
5'-AATTTTTgcCAACACCT-3'	*mut*17
5'-TTTTTTTTATTTATGCAGAGG-3'	SV40 TATA
5'-GTTTCTATAAAAACGCC-3'	CA1 TATA
5'-AAGACTATAAAAGTGTC-3'	CaBP9k TATA

Fig. 6. A: EMSA using labelled *CA1* −118 bp to −102 bp (CPO17) as probe and HT115 protein extracts. Each competitor oligonucleotide was at 100-fold molar excess of the probe. Comp = competitor; − = no competitor. 9k = *CaBP9k* TATA; SV40 = SV40 TATA and CA1 = the *CA1* TATA. COF1/COF1′ are complexes formed between the probe and the intestinal specific factor, Cdx2. B: Oligonucleotides used in the EMSA shown above.

there are several reports which show that in human colon carcinomas, *CA1* is markedly down-regulated [9,25].

The mechanism whereby *CA1* transcription is suppressed in the small intestine remains to be identified, but the presence of other regulatory sequences is indicated by the mapping of upstream DNase I hypersensitive sites as well as conserved protein binding DNA sequence motifs in the *CA1* colon promoter [12,26]. For example, conserved binding elements for the GATA family of transcription factors and hepatocyte nuclear factors (HNF-1) occur within the proximal promoter [26]; in other genes, such as epsilon-globin [27], serine dehydratase [28] and apolipoprotein B [29] such elements are known to function as negative *cis*-acting elements.

The possibility that Cdx2 can influence transcription by binding to the TATA box motif of the *CA1* gene promoter requires further investigation. This may be of significance since binding of Cdx2 to a TATA box motif in an intestinal gene promoter, calbindin 9K (*CaBP9k*) has been described and led to the proposal that Cdx2 might interact with other factors specific to intestinal cells to regulate intestine-specific expression via the TATA box [19].

Acknowledgements: The authors would like to thank Dr Peter Traber (University of Pennsylvania School of Medicine, Philadelphia, PA, USA) for supplying the Cdx2 expression vector.

References

[1] Swenson, E.R. (1991) in: Carbonic Anhydrases: Cellular Physiology and Molecular Genetics (Dodgson, S.J., Tashian, R.E., Gros, G. and Carter, N.D., Eds.), pp. 265–287, Plenum Press, New York.
[2] von Engelhardt, W., Gros, G., Burmester, M., Hansen, K., Becker, G. and Rechkemmer, G. (1994) J. Physiol. 477, 365–371.
[3] Roediger, W.E.W. (1994) J. Parent. Ent. Nut. 18, 4–8.
[4] Lönnerholm, G., Selking, O. and Wistrand, P.J. (1985) Gastroenterology 88, 1151–1161.
[5] Parkkila, S., Parkilla, A.-K., Juvonen, T. and Rajaniemi, H. (1994) Gut 35, 646–650.
[6] Fleming, R.E., Parkkila, S., Parkkila, A.-K., Rajaniemi, H., Waheed, A. and Sly, W.S. (1995) J. Clin. Invest. 96, 2907–2913.
[7] Wang, Y., Harvey, C., Rousset, M. and Swallow, D.M. (1994) Pediatr. Res. 36, 514–521.
[8] Lönnerholm, G. and Wistrand, P. (1983) Biol. Neonate 44, 166–176.
[9] Sowden, J., Leigh, S., Talbot, I., Delhanty, J. and Edwards, Y. (1993) Differentiation 53, 67–74.
[10] Brady, H.J.M., Sowden, J.C., Edwards, M., Lowe, N. and Butterworth, P.H.W. (1989) FEBS Lett. 257, 451–456.
[11] Brady, H.J.M., Edwards, M., Linch, D.C., Knott, L., Barlow, J.H. and Butterworth, P.H.W. (1990) Br. J. Haematol. 76, 135–142.
[12] Drummond, F., Sowden, J., Morrison, K. and Edwards, Y.H. (1996) Eur. J. Biochem. 236, 670–681.
[13] Suh, E., Chen, L., Taylor, J. and Traber, P.G. (1994) Mol. Cell. Biol. 14, 7340–7351.
[14] Taylor, J.K., Boll, W., Levy, T., Suh, E.R., Siang, S., Mantei, N. and Traber, P.G. (1997) Gastroenterology 112, SS, A410.
[15] Troelsen, J.T., Mitchelmore, C., Spodsberg, N., Jensen, A.M., Norén, O. and Sjöström, H. (1997) Biochem. J. 322, 833–838.
[16] James, R. and Kazenwadel, J. (1991) J. Biol. Chem. 266, 3246–3251.
[17] James, R., Erler, T. and Kazenwadel, J. (1994) J. Biol. Chem. 269, 15229–15237.
[18] Gorman, C. (1985) in: DNA Cloning, Volume II. A Practical Approach (Glover, D.M., Ed.), pp. 143–190, IRL Press, Oxford.
[19] Lambert, M., Colnot, S., Suh, E.-R., L'Horset, L., Blin, C., Calliot, M.-L., Raymondjean, M., Thomasset, M., Traber, P.G. and Perret, C. (1996) Eur. J. Biochem. 236, 778–788.
[20] Jin, T. and Drucker, D.J. (1996) Mol. Cell. Biol. 16, 19–28.
[21] German, M.S., Wang, J., Chadwick, R.B. and Rutter, W.J. (1992) Genes Dev. 6, 2165–2176.
[22] Suh, E.-R. and Traber, P.G. (1996) Mol. Cell Biol. 16, 619–625.
[23] Taylor, J.K., Levy, T., Suh, E.-R. and Traber, P.G. (1997) Nucleic Acids Res. 25, 2293–2300.
[24] Chawengsaksophak, K., James, R., Hammond, V.E., Köntgen, F. and Beck, F. (1997) Nature 386, 84–87.
[25] Mori, M., Staniunas, R.J., Barnard, G.F., Jessup, J.M., Steele Jr., G.D. and Chen, L.B. (1993) Gastroenterology 105, 820–826.
[26] Drummond, F. (1997) Carbonic Anhydrase 1: A Study of Colon-specific Gene Regulation, Thesis.
[27] Raich, N., Clegg, C.H., Grofti, J., Romeo, P.H. and Stamatoyannopolous, G. (1995) EMBO J. 14, 801–80929.
[28] Noda, C., Fukushima, C., Fujiwara, T., Matsuda, K., Kobune, Y. and Ichihara, A. (1994) Biochim. Biophys. Acta 1217, 163–173.
[29] Lee, S.-Y., Nagy, B.P., Brooks, A.R., Wang, D.-M., Paulweber, B. and Levy-Wilson, B. (1996) J. Biol. Chem. 271, 707–718.